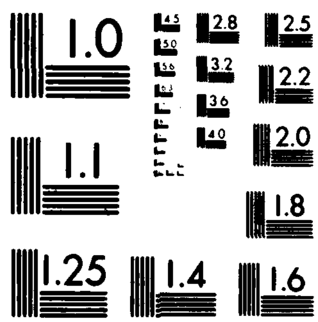
AD-A114 574 WISCONSIN UNIV-MADISON MATHEMATICS RESEARCH CENTER F/G 12/1
THE NUMERICAL SOLUTION OF A CLASS OF CONSTRAINED MINIMIZATION P--ETC(U)
FEB 82 N DYN, W E FERGUSON DAAG29-80-C-0041
UNCLASSIFIED MRC-TSR-2340 NL

END
DATE
FILMED
6 82
DTIC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS 1963 A

ADA114574

MRC Technical Summary Report #2340

THE NUMERICAL SOLUTION OF A CLASS OF CONSTRAINED MINIMIZATION PROBLEMS

Nira Dyn and Warren E. Ferguson, Jr.

Mathematics Research Center
University of Wisconsin—Madison
610 Walnut Street
Madison, Wisconsin 53706

February 1982

(Received October 23, 1981)

Sponsored by

U. S. Army Research Office
P. O. Box 12211
Research Triangle Park
North Carolina 27709

UNIVERSITY OF WISCONSIN - MADISON
MATHEMATICS RESEARCH CENTER

THE NUMERICAL SOLUTION OF A CLASS OF CONSTRAINED MINIMIZATION PROBLEMS

Nira Dyn* and Warren E. Ferguson, Jr.**

Technical Summary Report #2340

February 1982

ABSTRACT

This paper proves that a large class of iterative schemes can be used to solve a certain constrained minimization problem. The constrained minimization problem considered involves the minimization of a quadratic functional subject to linear equality constraints. Among this class of convergent iterative schemes are generalizations of the relaxed Jacobi, Gauss-Seidel, and symmetric Gauss-Seidel schemes.

AMS (MOS) Subject Classification: 49D40, 65F10, 65N20

Key Words: Constrained minimization, Quadratic programming, Iterative schemes, Indefinite systems of equations.

Work Unit Number 3 - Numerical Analysis and Computer Science

* School of Mathematical Sciences, Tel-Aviv University, Israel.

** Department of Mathematics, Southern Methodist University, Dallas, Texas 75275.

Sponsored by the United States Army under Contract No. DAAG29-80-C-0041.

SIGNIFICANCE AND EXPLANATION

Consider the following problem: Find the real n-vector x_* which minimizes $f(x) \equiv \frac{1}{2}x^TAx - x^Tr$ subject to the constraints $g(x) \equiv E^Tx - s = 0$. The purpose of this paper is to show that there is a large class of iterative schemes which can be used to solve such problems. These schemes are particularly effective when the matrix A is large and sparse and there are only a moderate number of constraints; one example of such a problem is described at the end of this paper.

The responsibility for the wording and views expressed in this descriptive summary lies with MRC, and not with the authors of this report.

THE NUMERICAL SOLUTION OF A CLASS OF CONSTRAINED MINIMIZATION PROBLEMS

Nira Dyn* and Warren E. Ferguson**

§1. Introduction

In this paper we will present several iterative schemes which solve the following constrained minimization problem:

Problem 1: Find the real n-vector x_* which minimizes $f(x) \equiv \frac{1}{2} x^T Ax - x^T r$

subject to the constraints $g(x) \equiv E^T x - s = 0$.

Here A is a real symmetric nonnegative definite $n \times n$ matrix, E is a real $n \times m$ matrix with full column rank, r is a real n-vector, and s is a real m-vector.

As discussed in section 2, the theory of quadratic programming [Hadley] states that under reasonable conditions on A and E the solution of Problem 1 exists and is unique. Furthermore, under these conditions on A and E the solution x_* of Problem 1 is the x component of the solution (x_*,λ_*) of the following problem:

Problem 2: Find the real n-vector x_* and the real m-vector λ_* which

solves the linear system

$$\begin{bmatrix} A & E \\ E^T & 0 \end{bmatrix} \begin{bmatrix} x \\ \lambda \end{bmatrix} = \begin{bmatrix} r \\ s \end{bmatrix} .$$

In section 3 we describe the convergence of a large class of iterative schemes used to solve Problem 2, and hence Problem 1. Although our iterative schemes are generally applicable to these problems, they are typically efficient only when A is a large sparse matrix and there are only a moderate number of constraints. In this situation the usual methods used to solve these problems become inefficient. The application of similar iterative schemes to the minimization of a quadratic form under inequality constraints is investigated in [Mangasarian].

* School of Mathematical Sciences, Tel-Aviv University, Israel.

** Department of Mathematics, Southern Methodist University, Dallas, Texas 75275.

Sponsored by the United States Army under Contract No. DAAG29-80-C-0041.

Our work was motivated by the work of [Tobler] in which a variant of one of the iterative schemes described in this paper was used to numerically construct a smooth surface from aggregated data. This application is described in section 4. The numerical solution of Problems 1 and 2 has also been considered by other authors; in particular we mention the work presented by [Luenberger], [Paige, Saunders], and [Gill, Murray].

§2. Preliminaries

In the previous section we stated that under reasonable conditions on A and E the solution x_* of Problem 1 is part of the solution (x_*,λ_*) of Problem 2. This statement is contained in the following theorem.

Theorem 2.1: Assume that

(a) A is a real symmetric nonnegative definite matrix,

(b) E is a real matrix with full column rank, and

(c) A and E^T have no nontrivial null vectors in common.

Then the solutions of Problems 1 and 2 exist and are unique. Furthermore: if x_* is the solution of Problem 1 then $(x_*, (E^TE)^{-1}E^T(r - Ax_*))$ is the solution of Problem 2, if (x_*,λ_*) is the solution of Problem 2 then x_* is the solution of Problem 1.

Proof: See the treatment of quadratic programming given in [Hadley]. ∎

Corollary 2.2: Under the assumptions of Theorem 2.1 the matrix

$$\begin{bmatrix} A & E \\ E^T & 0 \end{bmatrix}$$

is nonsingular.

Proof: Observe that this matrix is the coefficient matrix of the linear system in Problem 2. Under the assumptions of Theorem 2.1 this linear system has only unique solutions, therefore as shown in [Stewart] the coefficient matrix is nonsingular. ∎

The iterative schemes we use to solve Problem 2 are all based upon a splitting

$$A = B - C$$

of the matrix A, and they have the following form. Given an initial iterate (x_0,λ_0) define (x_k,λ_k) for $k = 1,2,3,\ldots$ to be the solution of the linear system

$$\begin{bmatrix} B & E \\ E^T & 0 \end{bmatrix} \begin{bmatrix} x_k \\ \lambda_k \end{bmatrix} = \begin{bmatrix} C & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_{k-1} \\ \lambda_{k-1} \end{bmatrix} + \begin{bmatrix} r \\ s \end{bmatrix} . \qquad (1)$$

Of course, for this linear stationary iterative method to be well defined it is necessary and sufficient that the matrix

$$\begin{bmatrix} B & E \\ E^T & 0 \end{bmatrix}$$

be nonsingular. This problem is addressed by the following theorem.

Theorem 2.3: In addition to the assumptions of Theorem 2.1 let

(d) $A = B - C$,

(e) B be a real nonsingular matrix, and

(f) $2A + C + C^T$ be a positive definite matrix.

Then the iterative scheme (1) is well-defined.

Proof: Since B is nonsingular then

$$\begin{bmatrix} B & E \\ E^T & 0 \end{bmatrix} = \begin{bmatrix} B & 0 \\ E^T & -E^T B^{-1} E \end{bmatrix} \begin{bmatrix} I & B^{-1}E \\ 0 & I \end{bmatrix} ,$$

therefore it follows that the matrix on the left hand side of the equality is nonsingular if and only if $E^T B^{-1} E$ is nonsingular. To prove that $E^T B^{-1} E$ is nonsingular let us prove that $E^T B^{-1} E$ has no nontrivial null vector. If $E^T B^{-1} E\lambda = 0$ then

$$\begin{aligned} 0 &= \lambda^T E^T B^{-1} E\lambda \\ &= (B^{-1}E\lambda)^T B^T (B^{-1}E\lambda) \\ &= \tfrac{1}{2} (B^{-1}E\lambda)^T (B + B^T)(B^{-1}E\lambda) \\ &= \tfrac{1}{2} (B^{-1}E\lambda)^T (2A + C + C^T)(B^{-1}E\lambda) \quad . \end{aligned}$$

This implies $B^{-1}E\lambda = 0$, since $2A + C + C^T$ is a positive definite matrix, and so $\lambda = 0$ because E is a matrix with full column rank.

∎

Let us now describe one procedure for solving the linear system (1) for (x_k, λ_k).

Step 1: Solve

$$B\hat{x}_k = Cx_{k-1} + r$$

for $\hat{x}_k$.

Step 2: Solve

$$(E^T B^{-1} E)\lambda_k = E^T \hat{x}_k - s$$

for λ_k.

Step 3: Solve

$$B(x_k - \hat{x}_k) = -E\lambda_k$$

for $x_k - \hat{x}_k$.

Step 2 requires the solution of a full linear system of order m. Therefore the iterative scheme is efficient only when $m \ll n$.

In the next section we will see that if assumption (f) of Theorem 2.3 is slightly strengthened then the iterative scheme (1) is not only well-defined but also convergent.

§3. Convergence of the Iterative Schemes

One set of conditions which guarantees the convergence of the iterative scheme (1) is described in the following theorem.

Theorem 3.1: Assume that

(a) A is a real symmetric nonnegative definite matrix,

(b) E is a real matrix with full column rank,

(c) A and E^T have no nontrivial null vectors in common,

(d) $A = B - C$,

(e) B is a real nonsingular matrix, and

(f) $A + C + C^T$ is a positive definite matrix.

Then the iterative scheme (1) is well-defined and convergent.

Proof: From Theorem 2.1 we deduce that a solution of Problem 2 exists and is unique. Since A is a nonnegative definite matrix and $A + C + C^T$ is a positive definite matrix then $2A + C + C^T$ is a positive definite matrix. From Theorem 2.3 we therefore deduce that the iterative scheme (1) is well-defined.

As shown in [Wendroff], the iterative scheme (1) is convergent if and only if each eigenvalue of the matrix

$$\begin{bmatrix} B & E \\ E^T & 0 \end{bmatrix}^{-1} \begin{bmatrix} C & 0 \\ 0 & 0 \end{bmatrix} \tag{2}$$

has magnitude less than one. Let us therefore show that if μ is a nonzero eigenvalue of the matrix (2) then the magnitude of μ is less than one.

Since μ is an eigenvalue of the matrix (2) then there are complex vectors u,v not both zero for which

$$\mu \begin{bmatrix} B & E \\ E^T & 0 \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} C & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} . \tag{3}$$

Let us now argue that $u \neq 0$ and $u^H Au > 0$. Since A is a real symmetric nonnegative definite matrix then clearly $u^H Au \geq 0$, and $u^H Au = 0$ only if $Au = 0$.

However $Au = 0$ only if $u = 0$, for (3) states that $E^T u = 0$ and by hypothesis A and E^T have no nontrivial null vectors in common. But $u = 0$ only if $v = 0$, for by hypothesis E has full column rank and (3) implies that $Ev = 0$ when $u = 0$. Since u,v are not both zero then we conclude $u \neq 0$ and $u^H Au > 0$.

Let us now establish the fact that

$$(1 - |\mu|^2)u^H Au = |1 - \mu|^2 u^H (A + C + C^T)u \quad . \tag{4}$$

We begin with the identity

$$u^H Au - (\mu u)^H A(\mu u) = (u - \mu u)^H A(u - \mu u) + 2\mathrm{Re}\{(u - \mu u)^H A(\mu u)\} \quad . \tag{5}$$

Using (3), and the fact that $A = B - C$, we find that

$$\begin{aligned}(u - \mu u)^H A(\mu u) &= (u - \mu u)^H (B-C)(\mu u)\\ &= (u - \mu u)^H (\mu Bu - \mu Cu)\\ &= (u - \mu u)^H (Cu - \mu Ev - \mu Cu)\\ &= (u - \mu u)^H C(u - \mu u) \quad ,\end{aligned}$$

which reduces (5) to the result stated in (4).

By hypothesis, $A + C + C^T$ is a positive definite matrix. Since $u \neq 0$ we know that $u^H(A + C + C^T)u > 0$, and so (4) implies that either $|\mu| < 1$ or $\mu = 1$. However it is impossible that $\mu = 1$, for if $\mu = 1$ then (3) would imply that u,v are both zero because by Corollary 2.2 the matrix

$$\begin{bmatrix} A & E \\ E^T & 0 \end{bmatrix}$$

is nonsingular.

■

Let us now describe several iterative schemes whose convergence is assured by Theorem 3.1.

Corollary 3.2: Assume that

(a) A is a real symmetric nonnegative definite matrix,

(b) E is a real matrix with full column rank,

(c) A and E^T have no nontrivial null vectors in common,

(g) $A = D - L - L^T$ where D is a nonsingular diagonal matrix and L is a strictly lower triangular matrix.

Then the iterative scheme (1) is convergent for the following choices of B and C.

(1) $B = \frac{1}{\omega} D$, $C = (\frac{1-\omega}{\omega})D + L + L^T$ with $\omega > 0$ chosen so small that $\frac{2}{\omega} D - A$ is a positive definite matrix.

(2) $B = \frac{1}{\omega} D - L$, $C = \frac{1-\omega}{\omega} D + L^T$ with $0 < \omega < 2$.

(3) $B = (\frac{2-\omega}{\omega})^{-1}(\frac{1}{\omega} D - L)D^{-1}(\frac{1}{\omega} D - L)^T$,

$C = (\frac{2-\omega}{\omega})^{-1}(\frac{1-\omega}{\omega} D + L)D^{-1}(\frac{1-\omega}{\omega} D + L)^T$ with $0 < \omega < 2$.

Proof: It is clear that assumptions (a) thru (e) of Theorem 3.1 are valid for each of the above choices for B and C. Therefore the iterative scheme (1) will be convergent if assumption (f) of Theorem 3.1 is valid for each of the above choices of B and C. For the first choice of B and C we find that

$$A + C + C^T = \frac{2}{\omega} D - A$$

and so assumption (f) of Theorem 3.1 is valid if $\omega > 0$ is chosen so small that $\frac{2}{\omega} D - A$ is positive definite. For the second choice of B and C we find that

$$A + C + C^T = \frac{2-\omega}{\omega} D$$

and so assumption (f) of Theorem 3.1 is valid if $0 < \omega < 2$. For the third choice of B and C we note that

$$A + C + C^T = B + C \quad ,$$

where $B(C)$ is a real symmetric positive (non-negative) definite matrix if $0 < \omega < 2$, and so assumption (f) of Theorem 3.1 is valid if $0 < \omega < 2$.

■

We note that the first, second, and third choices of B and C described in Corollary 3.2 correspond respectively to the usual JOR, SOR, and SSOR splittings of A described in [Young]. Under further assumptions Corollary 3.2 can be extended to the line and block versions of JOR, SOR, and SSOR. Furthermore, there is an obvious generalization of Theorem 3.1 to complex matrices.

§4. Application of the Iterative Scheme

Let us consider the problem of estimating a certain geographically varying quantity over a bounded geographical region, given the integrals of the quantity over several disjoint sub-regions.

Let Ω represent this finite geographical region and $\Omega_1, \Omega_2, \ldots, \Omega_m$ its disjoint sub-regions. In addition let $v(x,y)$ represent the function over Ω whose values we wish to estimate.

We shall choose the function $u^*(x,y)$ by requiring that it minimize the functional

$$J(u) = \tfrac{1}{2} \int_\Omega |\nabla u(x,y)|^2 dxdy \tag{5}$$

among all functions in the Sobolev space $H^1(\Omega)$ which satisfy the constraints

$$\int_{\Omega_i} u(x,y) dxdy = s_i \quad \text{for} \quad i = 1,2,\ldots,m \quad . \tag{6}$$

Here the numbers $s_1, s_2, \ldots, s_m$ are the given values of the integral of $v(x,y)$ over the subregions $\Omega_1, \Omega_2, \ldots, \Omega_m$. This problem was considered by [Tobler] and analyzed in [Dyn, Wahba]; in particular the paper by [Dyn, Wahba] demonstrates that this problem is well-posed.

This problem is discretized by the finite element method as described in [Friedrichs, Keller]. Here a triangulation Ω_T of Ω is introduced and $u^*(x,y)$ is approximated by the continuous function $u(x,y)$ which is a linear function of x and y in each triangle of Ω_T, minimizes the functional (5), and satisfies the constraints (6).

Suppose that the vertices $P_1, P_2, \ldots, P_n$ of the triangles in Ω_T are labeled in some order. If the i-th component of the n-vector u represents the value of $u(x,y)$ at P_i then the functional (5) admits that representation

$$J(u) = \tfrac{1}{2} u^T A u$$

where A is a real symmetric nonnegative definite $n \times n$ matrix having positive diagonal entries. In addition the constraints (6) admit the representation

$$E^T u = s$$

where E is a real $n \times m$ matrix and s is a real m-vector whose i-th component is s_i. If the triangulation Ω_T is sufficiently fine then the matrix E will have full

column rank. Therefore the n-vector u can be determined by any of the iterative schemes described in Corollary 3.2.

The convergence of the solution u to the solution u^* of problem (5) and (6), as the triangulation become finer, is demonstrated in [Dyn, Ferguson].

REFERENCES

N. Dyn and W. Ferguson, Numerical Construction of Smooth Surfaces from Aggregated Data, University of Wisconsin-Madison, Mathematics Research Center, TSR #2129, October 1980.

N. Dyn and G. Wahba, On the estimation of functions of several variables from aggregated data, University of Wisconsin-Madison, Mathematics Research Center TSR #1974, July, 1979, to appear in SIAM J. Math. Anal.

K. O. Friedrichs and H. B. Keller, A finite difference scheme for generalized Neumann problems, in Numerical Solution of Partial Differential Equations, J. H. Bramble, ed., Academic Press, New York, New York (1966).

P. E. Gill and W. Murray, Numerical Methods for Constrained Optimization, Academic Press, New York, New York (1974).

G. Hadley, Nonlinear and Dynamic Programming, Addison-Wesley, Reading, Massachusetts (1964).

D. G. Luenberger, Hyperbolic Pairs in the Method of Conjugate Gradients, SIAM J. Appl. Math., 17 (1969), 1263-1267.

The Conjugate Residual Method for Constrained Minimization Problems, SIAM J. Num. Anal., 7 (1970), 390-398.

O. L. Mangasarian, Sparsity-Preserving SOR Algorithms for Separable Quadratic and Linear Programming, University of Wisconsin-Madison, Mathematics Research Center, TSR #2260, August 1981.

C. C. Paige and M. A. Saunders, Solution of Sparse Indefinite Systems of Equations, SIAM J. Num. Anal., 12 (1975), 617-629.

G. W. Stewart, Introduction to Matrix Computations, Academic Press, New York, New York (1973).

W. Tobler, Smooth Pycnophylactic Interpolation for Geographical Regions, J.A.S.A., 74 (1979), 519-530.

B. Wendroff, Theoretical Numerical Analysis, Academic Press, New York, New York (1966).

D. M. Young, Iterative Solution of Large Linear Systems, Academic Press, New York, New York (1971).

MD/WEF/jvs

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER #2340	2. GOVT ACCESSION NO. AD-A114 574	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) The Numerical Solution of a Class of Constrained Minimization Problems		5. TYPE OF REPORT & PERIOD COVERED Summary Report - no specific reporting period
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Nira Dyn and Warren E. Ferguson, Jr.		8. CONTRACT OR GRANT NUMBER(s) DAAG29-80-C-0041
9. PERFORMING ORGANIZATION NAME AND ADDRESS Mathematics Research Center, University of Wisconsin, 610 Walnut Street, Madison, Wisconsin 53706		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS Work Unit Number 3 - Numerical Analysis and Computer Science
11. CONTROLLING OFFICE NAME AND ADDRESS U. S. Army Research Office, P.O. Box 12211, Research Triangle Park, North Carolina 27709		12. REPORT DATE February 1982
		13. NUMBER OF PAGES 12
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Constrained minimization, Quadratic programming, Iterative schemes, Indefinite systems of equations.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This paper proves that a large class of iterative schemes can be used to solve a certain constrained minimization problem. The constrained minimization problem considered involves the minimization of a quadratic functional subject to linear equality constraints. Among this class of convergent iterative schemes are generalizations of the relaxed Jacobi, Gauss-Seidel, and symmetric Gauss-Seidel schemes.

DD FORM 1473, 1 JAN 73 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

DATE
FILME
-8